

Linköping Studies in Science and Technology
Dissertations, No. 1929

Optimizing Massive MIMO: Precoder Design and Power Allocation

Hei Victor Cheng



Division of Communication Systems
Department of Electrical Engineering (ISY)
Linköping University, 581 83 Linköping, Sweden
www.commsys.isy.liu.se

Linköping 2018

This is a Swedish Doctor of Philosophy thesis.
The Doctor of Philosophy degree comprises 240 ECTS credits of
postgraduate studies.

**Optimizing Massive MIMO: Precoder Design and Power
Allocation**

© 2018 Hei Victor Cheng, unless otherwise noted.

ISBN 978-91-7685-327-6

ISSN 0345-7524

Printed in Sweden by LiU-Tryck, Linköping 2018

感谢我的父母,

&

献给我最爱的玥华!

Abstract

The past decades have seen a rapid growth of mobile data traffic, both in terms of connected devices and data rate. To satisfy the ever growing data traffic demand in wireless communication systems, the current cellular systems have to be redesigned to increase both spectral efficiency and energy efficiency. Massive MIMO (Multiple-Input-Multiple-Output) is one solution that satisfy both requirements. In massive MIMO systems, hundreds of antennas are employed at the base station to provide service to many users at the same time and frequency. This enables the system to serve the users with uniformly good quality of service simultaneously, with low-cost hardware and without using extra bandwidth and energy. To achieve this, proper resource allocation is needed. Among the available resources, transmit power beamforming are the most important degrees of freedom to control the spectral efficiency and energy efficiency. Due to the use of excessive number of antennas and low-end hardware at the base station, new aspects of power allocation and beamforming compared to current systems arises.

In the first part of the thesis, new uplink power allocation schemes that based on long term channel statistics is proposed. Since quality of the channel estimates is crucial in massive MIMO, in addition to data power allocation, joint power allocation that includes the pilot power as additional variable should be considered. Therefore a new framework for power allocation that matches practical systems is developed, as the methods developed in the literature cannot be applied directly to massive MIMO systems. Simulation results confirm the advantages brought by the the proposed new framework.

In the second part, we introduces a new approach to solve the joint precoding and power allocation for different objective in downlink scenarios by a combination of random matrix theory and optimization theory. The new approach results in a simplified problem that, though non-convex, obeys a

simple separable structure. Simulation results showed that the proposed scheme provides large gains over heuristic solutions when the number of users in the cell is large, which is suitable for applying in massive MIMO systems.

In the third part we investigate the effects of using low-end amplifiers at the base stations. The non-linear behavior of power consumption in these amplifiers changes the power consumption model at the base station, thereby changes the power allocation and beamforming design. Different scenarios are investigated and results show that a certain number of antennas can be turned off in some scenarios.

In the last part we consider the use of non-orthogonal-multiple-access (NOMA) inside massive MIMO systems in practical scenarios where channel state information (CSI) is acquired through pilot signaling. Achievable rate analysis is carried out for different pilot signaling schemes including both uplink and downlink pilots. Numerical results show that when downlink CSI is available at the users, our proposed NOMA scheme outperforms orthogonal schemes. However with more groups of users present in the cell, it is preferable to use multi-user beamforming in stead of NOMA.

Populärvetenskap

Sammanfattning

De senaste årtiondena har den mobila datatrafiken ökat snabbt, både i termer av antal uppkopplade enheter och mängd data. För att tillfredsställa den ständigt växande efterfrågan på trådlös kommunikation måste de nuvarande cellulära systemen omdanas för att öka både deras spektraleffektivitet och energieffektivitet. Massiv MIMO (eng: multiple-input-multiple-output) är en lösning som uppfyller båda kraven. I massiv MIMO används hundratals antenner vid basstationen för att betjäna många användare över samma tid-frekvensresurs. Detta gör att systemet simultant kan betjäna alla användare med samma goda betjäningsgrad, med enkel hårdvara och utan att använda större bandbredd eller mer energi. För att uppnå detta behöver systemets resurser allokeras korrekt. Bland de tillgängliga resurserna är lobformning och effektregering de viktigaste frihetsgraderna att allokera för att förbättra spektraltätheten och energieffektiviteten. På grund av det stora antalet antenner och den enkla hårdvaran vid basstationen dyker nya aspekter hos effektallokeringen och lobformningen upp jämfört med nuvarande system.

I avhandlingens första artikel föreslås nya effektallokeringsmetoder för upplänken, vilka är baserade på den långsiktiga kanalstatistiken. Eftersom kvaliteten hos kanalestimaten är viktig i massiv MIMO, bör även effekten hos piloterna regleras, inte bara effekten hos den databärande signalen. Ett nytt ramverk för sådan simultan effektregering utvecklas, eftersom tidigare metoder i litteraturen inte direkt kan appliceras på massiv MIMO. Simuleringar bekräftar fördelarna med det föreslagna, nya ramverket.

I den andra artikeln introduceras ett nytt sätt att simultant optimera förkodning och effektallokering för olika målfunktioner i nerlänken, genom en kombination av slumpmatristeori och optimering. Den nya metoden resulterar i ett förenklat problem som, ehuru icke-konvext, har en enkel separabel struktur. Simuleringsresultat visar att den föreslagna metoden ger stora vinster

över heuristiska lösningar när antalet användare i cellen är stort, vilket ofta är fallet i massiv MIMO. I den tredje och fjärde artikeln undersöker vi effekten av att använda enkla förstärkare vid basstationen. Det icke-linjära sättet effektförbrukningen uppför sig på hos dessa förstärkare ändrar effektförbrukningsmodellen vid basstationen, och därigenom ändras även utformningen av effektallokeringen och lobformningen. Olika scenarier undersöks och resultaten visar att ett givet antal antenner kan stängas av i vissa scenarier.

I den femte artikeln betraktar vi användandet av NOMA (icke-ortogonal fleranvändaråtkomst, eng: non-orthogonal multiple access) i massiv MIMO i praktiska scenarier, där kanalkänedom erhålls genom pilotsignaler. Datatoken analyseras för olika pilotsignaleringsmetoder, med både upplänks- och nerlänkspiloter. Numeriska resultat visar att, när kanalkänedom om nerlänken finns tillgänglig hos användarna, presterar vår föreslagna NOMA-metod bättre än ortogonala åtkomstmetoder. Det är emellertid bättre att använda fleranvändarlobformning än NOMA när flera grupper av användare är närvarande i cellen.

Acknowledgments

I would like to express my greatest gratitude to my supervisor, Prof. Erik G. Larsson, for offering me the opportunity to study for the PhD degree in his group. He provided me with guidance and advice on research problems. Discussions with Erik are always exciting and challenging. His insightful comments on my work helped me to develop new ideas and find the right way to proceed. Apart from technical aspects, I also learnt from him the structured way of working and the ability to handle multiple tasks simultaneously. I am sure his perseverance in the pursuit of top-quality research will have a deep impact on my future life.

I would also like to thank Dr. Daniel Persson, now at Qamcom Research and Technology, Gothenburg, Sweden, who has been my co-supervisor during my studies in the first two years. He helped me to adapt to a new environment when I just join the group. I had learnt a lot from him, especially in the area of amplifiers and video coding. I am also grateful to Prof. Emil Björnson, who is my current co-supervisor. He is always there to help whenever I face problems in my research. His rich knowledge on optimization solves many of my doubts and curiosities. Further, he helped me a lot on improving my writing style which is particularly appreciated. Finally I would like to thank all members in the research group Communication Systems for all the interesting discussions and chats.

Finally, I would like to dedicate my sincere thanks to my parents and my dearest girlfriend Yuehua for their unconditional love and support in my life.

Hei Victor Cheng
Linköping, March 2018

Contents

Abstract	v
Populärvetenskap Sammanfattning	vii
Acknowledgments	ix
1 Introduction	1
1.1 Background	1
1.2 Contributions of the thesis	5
1.3 Papers Not Included in the Thesis	8
2 Power Allocation in Wireless Networks	11
2.1 From Orthogonal Access to Power Allocation	12
2.2 Power Allocation with Target SINR	12
2.3 Power Allocation with Different Objectives	15
2.3.1 Weighted Max-Min Fairness	16
2.3.2 Weighted Sum Performance	16
2.4 Joint Beamformer Design and Power Allocation	17
2.5 Receiver with Successive Interference Cancellation Capability	18
3 Power Allocation and Precoder Design in Massive MIMO	21
3.1 The Role of CSI	22
3.2 Joint Pilot and Data Power Allocation	23
3.3 Help of Asymptotic Analysis	24
3.4 Power Allocation with Imperfect Amplifiers	25
4 Background on Optimization	27
4.1 General Optimization Problems	27
4.2 Some Convexity Preserving Transformation	28
4.2.1 Epigraph	29
4.2.2 Variable Substitution	29

4.2.3	Perspective Transformation	30
4.3	General Sufficient Conditions for Global Optimality	31
4.4	Successive Convex Optimization	33
5	Future Work	35
	Bibliography	36
	Included Papers	45
A	Optimal Pilot and Payload Power Control	47
1	Introduction	49
1.1	Background and Motivation	49
1.2	Related Work and Our Contributions	51
2	System Model	54
2.1	Achievable SE With Linear Detection	55
3	Optimal Pilot Length	57
4	Joint Power Control for Weighted Max-Min SE	58
4.1	Max-Min for MRC	58
4.2	Max-Min for ZF	63
5	Joint Pilot and Data Power Control for Weighted Sum SE	64
5.1	Weighted Sum SE for MRC	65
5.2	Sum SE for ZF	70
6	Extension to Correlated Fading Channels	71
7	Simulation Results and Discussion	73
7.1	Max-Min SE Results	74
7.2	Sum SE Results	77
7.3	Robustness	80
7.4	Correlated Fading	81
7.5	Dependence on SNR, K and T	82
7.6	Complexity	84
8	Conclusion	85
	Appendix A: Proof of Theorem 1	86
	Appendix B: Proof of Theorem 5	88
	Appendix C: Proof of Lemma 6	89
	References	90
B	Optimal Precoder Design Via Large-System Analysis	95
1	Introduction	97
1.1	Related Work	98

1.2	Contributions of this Paper	98
2	System Model	99
3	Optimal Linear Precoding	100
3.1	Optimal Precoding for Power Minimization with SINR Targets	100
3.2	Optimal Precoder for General Utility Maximization	102
4	Sufficient and Necessary Optimality Conditions	103
5	Examples of Separable Utility Functions	105
5.1	Mean Square Error Minimization	105
5.2	Sum Rate Maximization	106
5.3	Discussion	107
6	Extensions	108
6.1	Max-Min Fairness Maximization	108
6.2	Imperfect CSI at the BS	109
7	Simulation Results	111
8	Conclusion	114
	Appendix A: Proof of Lemma 1	115
	Appendix B: Proof of Theorem 1	115
	Appendix C: Proof of Theorem 2	116
	Appendix D: Proof of Proposition 1	118
	References	119

C Precoding Under Amplifier Power Consumption Constraint **123**

1	Introduction	125
1.1	Technical contributions of this work	127
2	Capacity of Point to Point MIMO Channels	128
3	Upper Bound and Lower Bound on Capacity	130
3.1	Lower Bound on Capacity	130
3.2	Upper Bound on Capacity	131
3.3	Efficient Algorithm for Computation of a Lower Bound	132
4	Monotonic Optimization Approach for Global Optimal Solution	134
4.1	Branch and Bound Algorithm	134
4.2	Pruning to Speedup	136
4.3	Numerical Example of Global Optimization	136
5	Extension to Frequency-selective Point-to-point MIMO	138
6	Extension to Multiuser MIMO	140
7	Numerical Experiments	142
7.1	Schemes Included in the Comparison	143
7.2	Simulation Results	143
7.3	Discussion	149

8	Conclusion	150
	Appendix A: Proof of Statements	150
	References	151
D	Massive MIMO at Night	155
1	Introduction	157
2	Notations	159
3	Background	160
4	Power Minimization at Low System Load	162
5	Solution to the Power Minimization Problem	163
6	Extensions to Include Other Circuit Power Consumptions	165
7	Numerical Results	168
8	Conclusions	168
	References	172
E	NOMA in Training Based Multiuser MIMO Systems	175
1	Introduction	177
2	System Model	179
	2.1 Orthogonal Access Scheme	180
	2.2 Proposed NOMA Scheme	181
3	Uplink Channel Estimation	182
	3.1 MMSE Channel Estimation for Scheme-O	184
	3.2 MMSE Channel Estimation for Scheme-N	184
	3.3 Interference-Limited Scenarios	186
4	Performance Analysis	186
	4.1 Downlink Signal Model	187
	4.2 Performance With Perfect CSI at the Users	189
	4.3 Performance Without Downlink CSI	191
	4.4 Performance With Estimated Downlink CSI	196
5	Practical Issues and Extensions	200
	5.1 User Pairing	200
	5.2 More than Two Users Per Group	201
	5.3 Users with Multiple Antennas	201
	5.4 Power Control	201
6	Other Applications	202
	6.1 Application in Multicasting	202
	6.2 Rate-Splitting for Improving Sum Degree of Freedom	202
7	Numerical Results	203
	7.1 Small-Scale Antenna Systems	203
	7.2 Constrained Sum Rate Comparison	208

7.3	Effect of Number of Users or Number of Antennas at the User	211
8	Conclusion	213
	Appendix A: Proof of Proposition 1	214
	Appendix B: Proof of Proposition 2	214
	References	215

Chapter 1

Introduction

1.1 Background

In the past twenty years, we have seen a tremendous growth in the demand for wireless data rate, and this trend is predicted to continue in the future [1]. Methods to satisfy the ever growing demand of data rates include cell densification, increasing bandwidth for transmission and increasing the spectral efficiency. Massive MIMO was proposed in [2] and has become one of the most promising ways to increase the spectral efficiency of wireless cellular systems, and is now being considered to be the key enabling technology for the 5th generation cellular network [3]. The idea of massive MIMO is to use a large number of antennas at the base station (BS) to serve multiple users in the same time and frequency resource block. Massive MIMO is usually operating in the time-division-duplex (TDD) mode, where channel reciprocity is utilized to obtain downlink channel estimates from the uplink channel estimates. The benefits of massive MIMO include but are not limited to the following:

1. Spectral efficiency (SE): massive MIMO can increase the spectral efficiency per cell as it can multiplex multiple users to obtain a multiplexing gain, which is the pre-log factor in the sum capacity formula.
2. Radiated energy efficiency (EE): massive MIMO offers a large array gain that is equal to the number of antennas at the BS. The array gain is the factor of increase in terms of the received signal power. Thus

to achieve the same data rate as in traditional systems, the transmit power in both uplink and downlink can be greatly reduced and thereby increase the energy efficiency.

3. Simpler processing: assuming a rich scattering channel, when the number of antennas increases, channel hardening occurs where small scale fading is averaged out. Moreover, channels between users become more and more orthogonal to each other. In such cases ($M \gg K$) simple linear processing that treats interference as noise is shown to perform close to the optimal capacity achieving schemes.
4. Scalable: when operating in TDD mode, the number of antennas can theoretically be scaled indefinitely because the channel estimation overhead is only limited by the number of users being served.

Other benefits include the ability to handle mobility, extending coverage range and robustness to low-precision hardware.

The performance analysis of massive MIMO is therefore an important research topic and has attracted much attention from different researchers, see for example [4, 5]. Both theoretical analysis [6–9] and hardware experiments [10, 11] have validated the benefits of massive MIMO mentioned above. In order to fully harvest the gain in spectral efficiency and energy efficiency, proper resource allocation is needed. This includes user scheduling, sub-carrier allocation, power allocation and precoder (receiver) design, which correspond to resources in the time, frequency, power and spatial domain. Among them, power allocation and precoder design has been the most important aspects to enhance the system's rate performance in the physical layer and it can be combined with frequency subcarrier allocation and user scheduling to boost the performance further.

Power allocation optimization in wireless networks has been an important research problem for decades, dating back to single-antenna wireless systems. Due to channel fading and the interference from other users, power allocation problems are usually hard to solve optimally, in particular NP-hardness was proven in [12] for the objective of maximizing the sum SE in single-antenna multicellular wireless networks in both uplink and downlink transmission, even with single-carrier transmission. For practical use, a reasonable approach is to develop algorithms achieving local optimality with affordable complexity, as done for example in [13].

Compared to power allocation in single-antenna systems, power allocation in massive MIMO networks is a relatively new topic. Accurate channel estimates are needed at the BS for carrying out coherent linear processing, e.g. uplink detection and downlink precoding. Due to the large number of antennas in massive MIMO, instantaneous channel knowledge, which is commonly assumed to be known perfectly in the power allocation literature, is hard to obtain. The literature on power allocation for multi-user MIMO, and even joint power allocation and beamformer optimization, see for example [14,15] and the references therein, did not consider the channel estimation error explicitly and the design criterion was based on instantaneous CSI. Therefore the power allocation needs to be recalculated very frequently, and we denote it as short-term power allocation. Wireless channels only remain approximately constant over a certain coherence time and coherence bandwidth. The total number of symbols where the channels stay constant is called the coherence interval in the massive MIMO literature. For short-term power allocation, the parameters need to be calculated whenever the channels change (both over time and frequency). Therefore short-term power allocation is not practical as it introduces significant computational and communication overheads, which comes from the control signaling for adjusting the transmit power.

To overcome the disadvantages of short-term power allocation, we want to provide power allocation schemes that optimize the ergodic SE based on only the large-scale fading and we call this long-term power allocation. The ergodic SE is a commonly used metric to characterize the performance when CSI is not available at the transmitter and it assumes that the small-scale fading changes through all possible states of a ergodic stationary random process. Therefore a particular realization of the small-scale fading does not affect the ergodic SE. Due to the channel hardening effects in massive MIMO, the gain from short-term power allocation over long-term power allocation in terms of the ergodic SE is marginal. With long-term power allocation the same power allocation can be applied to different frequency sub-carriers and remain unchanged for a longer time because the large scale fading remains the same for a long time and over a wider bandwidth. Long-term power allocation reduces the computational overhead at the BS and the control signaling overhead, it also greatly simplifies the system design as the sub-carrier allocation problem becomes much simpler. Since the channel estimates are crucial in massive MIMO to perform beamforming, we also take into account the channel estimation errors in the power allocation model which introduces the pilot power as an additional optimization variable. Therefore, in this the-

sis, we develop a new framework for power allocation that matches practical systems (i.e., ergodic SE and imperfect CSI), as the methods developed in the literature cannot be applied directly for massive MIMO systems. Moreover when the number of antennas grows, the hardware impairment effects at the BS cannot be neglected [16] as low-end hardware is often preferred when deploying the BSs to reduce the deployment cost. For example, the use of low-end high-power amplifiers used for transmission, changes the power consumption at the BS, which introduces new problems when we optimize the power allocation for the downlink.

In the following sections, we overview the evolution of topics in power allocation and introduce new aspects in power allocation that arise from taking into account the effects of imperfect channel estimation and imperfect amplifiers in the systems.

1.2 Contributions of the thesis

This thesis considers new aspects of power allocation brought by the use of massive MIMO system. The thesis consists four main contributions. The first one is the optimization and analysis of power allocation under a new framework inspired by the properties of massive MIMO systems. This new framework takes into account the channel estimation errors and the optimization is carried out only based on the long term channel statistics. The second one combine the large-system analysis with optimization tools to tackle the precoder design in massive MIMO in a novel way. The third and fourth investigate the effects of using low-end amplifiers at the base stations. The non-linear behavior of power consumption in these amplifiers changes the power consumption model at the base station, and thereby changes the power allocation. The fifth analyses the performance of NOMA in multiuser MIMO when taking into account training overheads and channel estimation errors.

Brief summaries of the papers included in this thesis are as follows:

Paper A: Optimal Pilot and Payload Power Control in Single-Cell Massive MIMO Systems

Authored by: Hei Victor Cheng, Emil Björnson, and Erik G. Larsson

Published in: IEEE Transactions on Signal Processing, vol. 65, no. 9, pp. 2363-2678, May 2017 [17]. This work is an extension of the conference paper [18].

This paper considers the jointly optimal pilot and data power allocation in single-cell uplink massive multiple-input-multiple-output systems. Using the spectral efficiency as performance metric and setting a total energy budget per coherence interval, the power control is formulated as optimization problems for two different objective functions: the weighted minimum SE among the users and the weighted sum SE. A closed form solution for the optimal length of the pilot sequence is derived. The optimal power control policy for the former problem is found by solving a simple equation with a single variable. Utilizing the special structure arising from imperfect channel estimation, a convex reformulation is found to solve the latter problem to global optimality in polynomial time. The gain of the optimal joint power control is theoretically justified, and is proved to be large in the low SNR regime. Simulation results also show the advantage of optimizing the power control

over both pilot and data power, as compared to the cases of using full power and of only optimizing the data powers as done in previous work.

Paper B: Optimal Precoder Design in Downlink Multiuser MIMO Via Large-System Analysis

Authored by: Hei Victor Cheng, Emil Björnson, and Erik G. Larsson

Under preparation for submission to IEEE Transactions on Signal Processing. Part of this work will be presented in the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2018 [19].

This work introduces a new approach to find the solution to the joint precoding and power allocation problem in downlink multiuser MIMO, exploiting a combination of random matrix theory and optimization theory. The new approach results in simplified problems that, though non-convex, obey a simple separable structure. Different utilities are considered in the optimization problem, namely max-min fairness, sum mean-square error (MSE) and sum rate. The multi-variable optimization problems are decomposed into multiple single-variable optimization problems that can be solved in parallel. For the max-min fairness problem, a closed-form solution is found. For the sum MSE problem and the sum rate problem, water-filling-like solutions are found. The proposed approach is extended to handle the case of imperfect channel state information (CSI). The proposed approach provides large gains over heuristic solutions with similar computational complexity. The gain is increasing with the number of users in the cell, both in perfect and imperfect CSI scenarios, which suggests the applicability in massive MIMO systems.

Paper C: Optimal MIMO Precoding Under a Constraint on the Amplifier Power Consumption

Authored by: Hei Victor Cheng, Daniel Persson, and Erik G. Larsson

Submitted to IEEE Transactions on Communications, Dec. 2017. Part of this work was published in the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2014 [20].

The capacity of the MIMO channel taking into account both a limitation on total consumed power, and per-antenna radiated power constraints is considered. The total consumed power takes into account the traditionally used sum radiated power, and also the power dissipation in the amplifiers.

For a fixed channel with full CSI at both the transmitter and the receiver, maximization of the mutual information is formulated as an optimization problem. Lower and upper bounds on the capacity are provided by numerical algorithms based on partitioning of the feasible region. Both bounds are shown to converge and give the exact capacity when number of regions increases. The bounds are also used to construct a monotonic optimization algorithm based on the branch-and-bound approach. An efficient suboptimal algorithm based on successive convex approximation performing close to the capacity is also presented. Numerical results show that the performance of the solution obtained from the suboptimal algorithm is close to that of the global optimal solution. Simulation results also show that in the low SNR regime, antenna selection is the optimal scheme while at high SNR uniform power allocation is close to optimal.

Paper D: Massive MIMO at night: On the operation of massive MIMO in low traffic scenarios

Authored by: Hei Victor Cheng, D. Persson, Emil Björnson and Erik G. Larsson

Published in Proceeding of IEEE International Conference on Communications (ICC), pp. 1697 - 1702, 2015 [21].

For both maximum ratio transmission (MRT) and zero forcing (ZF) precoding schemes and given any specific rate requirement the optimal transmit power, number of antennas to be used, number of users to be served and number of pilots spent on channel training are found with the objective to minimize the total consumed power at the base station. The optimization problem is solved by finding closed form expressions of the optimal transmit power and then search over the remaining discrete variables. The analysis consists of two parts, the first part investigates the situation when only power consumed in the RF amplifiers is considered. The second part includes both the power consumed in the RF amplifiers and in other transceiver circuits. In the former case having all antennas active while reducing the transmit power is optimal. Adaptive scheme to switch off some of the antennas at the base stations is found to be optimal in the latter case.

Paper E: Performance Analysis of NOMA in Training Based Multiuser MIMO Systems

Authored by: Hei Victor Cheng, Emil Björnson and Erik G. Larsson

Published in: IEEE Transactions on Wireless Communications, vol. 17, no. 1, pp. 372 - 385, Jan. 2018 [22]. Part of this work was published in the IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2017 [23].

This paper considers the use of non-orthogonal-multiple-access (NOMA) in multiuser MIMO systems in practical scenarios where channel state information (CSI) is acquired through pilot signaling. A new NOMA scheme that uses shared pilots is proposed. Achievable rate analysis is carried out for different pilot signaling schemes, including both uplink and downlink pilots. The achievable rate performance of the proposed NOMA scheme with shared pilot within each group is compared with the traditional orthogonal access scheme with orthogonal pilots. Our proposed scheme is a generalization of the orthogonal scheme, and can be reduced to the orthogonal scheme when appropriate power allocation parameters are chosen. Numerical results show that when downlink CSI is available at the users, our proposed NOMA scheme outperforms orthogonal schemes. However with more groups of users present in the cell, it is preferable to use multi-user beamforming instead of NOMA.

1.3 Papers Not Included in the Thesis

The following paper contains work done by the author but is not included in this thesis because they are either outside the scope of this thesis, or they are earlier versions of the papers included in the thesis.

1. H. V. Cheng, E. G. Larsson, "Some fundamental limits on frequency synchronization in massive MIMO," *Asilomar Conference on Signals, Systems and Computers*, pp. 1213–1217, 2013.
2. D. Kapetanovic, H. V. Cheng, W. H. Mow and F. Rusek, "Optimal two-dimensional lattices for precoding of linear channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2104–2113, May 2013.
3. H. V. Cheng, D. Persson and E. G. Larsson, "MIMO capacity under power

amplifiers consumed power and per-antenna radiated power constraints,” in *Proceedings of IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 179–183, 2014.

4. H. V. Cheng, E. Björnson and E. G. Larsson, “Uplink pilot and data power control for single cell massive MIMO systems with MRC,” in *Proceedings of International Symposium on Wireless Communication Systems (ISWCS)*, pp. 396–400, 2015.

5. D. Kapetanovic, H. V. Cheng, W. H. Mow and F. Rusek, “Lattice structures of precoders maximizing the minimum distance in linear channels,” *IEEE Transactions on Information Theory*, vol. 61, no. 2, pp. 908–916, Feb. 2015.

6. H. V. Cheng, E. Björnson and E. G. Larsson, “NOMA in multiuser MIMO systems with imperfect CSI,” in *Proceedings of IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, 2017.

7. H. V. Cheng, E. Björnson and E. G. Larsson, “Semi-closed form solution for sum rate maximization in downlink multiuser MIMO via large-system analysis,” in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, to appear, 2018.

Chapter 2

Power Allocation in Wireless Networks

In wireless communication networks, transmit power is one of the most important degrees of freedom to control the rate performance. The rate performance $R = B \times \text{SE}$ can be characterized by the SE and the allocated bandwidth. Due to the broadcast nature of wireless communications, interference is the limiting factor for most cellular communication links. To overcome this, properly controlling the transmit power from the BS and the mobile terminals is of vast importance. This problem has been extensively studied in the past, for both the uplink and downlink transmission. New methods and algorithms have been developed over the last decades. Here we review the evolution of power allocation from its original form to the new formulations we developed in this thesis. We consider mostly the uplink power allocation as there is a form of duality between the uplink and downlink problem. The only difference comes from the power constraints, in which uplink admits independent power constraints per user while the downlink requires a total power constraint per cell. The power constraints are introduced due to the fact that all amplifiers have limited emitted powers, and because of regulations by governmental agencies. Moreover the power consumption of the users is more critical than the power consumption at the BS as the battery is limited at the mobile terminals.

2.1 From Orthogonal Access to Power Allocation

In traditional wireless networks, both transmitters and receivers are equipped with only one antenna. In such case, power, spectrum and time are the only three resources to optimize over. To eliminate the effect of interference, one method is to apply orthogonal access. This includes frequency division multiple access (FDMA) and time division multiple access (TDMA). They can both be viewed as special cases of power allocation as the transmitter is allocating zero power to specific frequency bands or time slots. However, orthogonal access does not utilize the spectrum efficiently when the rate region is convex. Due to the scarce spectrum resources in cellular systems, orthogonal access to every user becomes inefficient as the number of users increases and every user will get a rather low rate. The possibility for multiple links to share the same spectrum at the same time is therefore important and then power allocation comes into the picture. When the number of users further increases, the interference becomes strong and orthogonal access is useful again by dividing the users into different groups and applying orthogonal access to eliminate inter-group interference. In such case, user scheduling and frequency subcarrier allocation are introduced and all three resources have to be jointly optimized. In this chapter our focus will be on the power allocation part.

One example of power allocation is implemented in the Global System for Mobile Communication (GSM) cellular networks, where power allocation is done by first estimating the signal-to-interference-plus-noise-ratio (SINR) of a user at the BS. Then the BS feeds back a quantized measurement to inform the user to increase, maintain or decrease the power level. This kind of scheme is simple but it does not provide the optimal performance as the power allocation scheme may not converge or converge to something undesired. Therefore rigorous treatment of power allocation is needed, and many researchers have discovered many exciting results.

2.2 Power Allocation with Target SINR

The first problem which is formulated in the literature [24] is the power minimization problem with target SINR constraints. The power allocation scheme in the GSM can be viewed as one example in this theme. The target SINRs are usually set according to the application. For example in GSM

the target SINRs are set to ensure that the quality of a voice call is within an acceptable range. Here we describe the problem formulation and some approaches to solve the problem. Denote the channel gain from K users to the BS normalized with the noise variance at the BS as h_1, \dots, h_K , and the power from the users as p_1, \dots, p_K . In the literature the channel gains are mostly assumed to be perfectly known at the BS and at the users.

Define the SINR at a terminal as

$$\text{SINR}_k = \frac{h_k p_k}{1 + \sum_{j \neq k} h_j p_j}. \quad (1)$$

Here the constant 1 comes from the normalized power of the additive noise. This SINR has a direct relation to the Shannon information theoretic SE when the interference is treated as noise, h_k is treated as deterministic constant and independent Gaussian signaling from the users is assumed, this relation can be written as:

$$R_k = \log_2(1 + \text{SINR}_k). \quad (2)$$

Note that when BS has the ability to perform interference cancellation, the above SINR is not valid. In that case assuming the decoding is perfect and the decoding order is $1, 2, \dots, K$, the modified SINR can be written as

$$\text{SINR}_k^{\text{SIC}} = \frac{h_k p_k}{1 + \sum_{j > k} h_j p_j}. \quad (3)$$

However the interference cancellation requires complicated signal processing and in practice the decoding is imperfect which leads to error propagation. We therefore focus on the case when interference is treated as noise. In this case, the problem of transmit power minimization with target SINR can be formulated as the following optimization problem:

$$\begin{aligned} & \underset{\{p_k\}}{\text{minimize}} && \sum_k p_k \\ & \text{subject to} && \text{SINR}_k \geq \lambda_k, \quad \forall k, \end{aligned} \quad (4)$$

where λ_k specifies the target SINR of user k .

This problem can be solved by using linear programming (LP) as the constraints can be modified to be linear in the power allocation variables. However this requires centralized processing, i.e. BS solves the problem and inform users their transmit power, which introduces a large communication

overhead to the system since short-term power allocation is performed in this case. As a result, distributed power allocation schemes are preferred and this is first proposed in [25] to solve the distributed power allocation with fixed point iteration. This is done by rewriting the problem (4) in a matrix form:

$$\begin{aligned} & \underset{\mathbf{p}}{\text{minimize}} && \mathbf{1}^T \mathbf{p} \\ & \text{subject to} && (\mathbf{I} - \mathbf{G})\mathbf{p} \geq \mathbf{v} \end{aligned} \quad (5)$$

with the elements of \mathbf{G} being

$$G_{ij} = \begin{cases} \gamma_i \frac{h_j}{h_i}, & i \neq j \\ 0, & i = j \end{cases} \quad (6)$$

and the elements of \mathbf{v} being

$$v_i = \frac{\lambda_i}{h_i}. \quad (7)$$

Then by observing the fact that all SINR constraints will be achieved with equality at the optimal solution. The optimal power allocation can be found by solving the equation

$$\mathbf{p} = (\mathbf{I} - \mathbf{G})^{-1} \mathbf{v}. \quad (8)$$

This equation will have a unique positive solution (all elements in \mathbf{p} are positive) if and only if the spectral radius $\rho(\mathbf{G})$ of matrix \mathbf{G} is smaller than 1, i.e. $\rho(\mathbf{G}) < 1$. When this condition holds, the solution can be found from a fixed point iteration

$$\mathbf{p}(l+1) = \mathbf{G}\mathbf{p}(l) + \mathbf{v}, \quad (9)$$

or written in a more illustrative form for each user i

$$p_i(l+1) = \frac{\gamma_i}{\text{SINR}_i(l)} p_i(l). \quad (10)$$

From (10) a distributed power allocation scheme is established, where each user computes its own SINR and updates its transmit power according to the computed SINR. Later in [26] a more general framework is established. The author showed that when the general interference function $I(\mathbf{p})$ used in the fixed point equation is standard, the fixed point iteration method with $\mathbf{p}(k+1) = I(\mathbf{p}(k))$ will always converge to the global optimal solution with any initialization. The interference function can be arbitrary and, in the case we discussed above, it is $\mathbf{G}\mathbf{p}(k) + \mathbf{v}$.

The standard interference function is a function that has the following properties:

1. Positivity: $I(\mathbf{p}) \geq 0$.
2. Monotonicity: if $\mathbf{p} \geq \mathbf{q}$, then $I(\mathbf{p}) \geq I(\mathbf{q})$.
3. Scalability: for any $c > 1$, $cI(\mathbf{p}) > I(c\mathbf{p})$.

With this framework, modified schemes can be applied for different purposes. For example one scheme is to reduce the fluctuations in users' transmit power by taking

$$p_i(k+1) = \alpha p_i(k) + (1 - \alpha) \frac{\gamma_i}{\text{SINR}_i(k)} p_i(k). \quad (11)$$

2.3 Power Allocation with Different Objectives

Although the power minimization problem is easy to solve and admits a distributed solution, it is not an easy task to determine the target SINRs. This is because the wireless networks have evolved from providing voice call service to also providing mobile data. Different applications have different target data rate and it is no longer easy to set a target data rate such that every user will be satisfied. Moreover if the SINRs are not assigned properly, the data rates of the users may not be Pareto optimal or even infeasible. When the target SINRs are not feasible, the algorithm that were mentioned above will not converge. Therefore a better way to formulate the problem is to find the SINRs and the power allocation simultaneously. As the SINRs and SE are interchangeable, we formulate the general problem in terms of the data rate as follows:

$$\begin{aligned} & \underset{\{p_k\}}{\text{maximize}} && U(R_1, \dots, R_K) \\ & \text{subject to} && p_k \leq P_k, \forall k, \\ & && p_k \geq 0, \forall k, \end{aligned} \quad (12)$$

where the utility function $U(R_1, \dots, R_K)$ can be any kind of target function that is monotonically non-decreasing in every component, and P_k represents the power constraint at the user. Depending on the utility function we choose, the solution to (12) can be vastly different. Some of them are easy to solve and some are harder. With proper choice of utility function, we can achieve different points on the Pareto optimal boundary. Which means we cannot increase the rate of any of the users R_k without lowering the rate

of the other users. In the following we introduce two most commonly used utility functions, namely weighted max-min fairness problem and weighted sum problem.

2.3.1 Weighted Max-Min Fairness

The weighted max-min fairness problem is used to provide the same quality-of-service to all users in the cell. With max-min fairness we aim at serving every user with equal weighted SE according to their priorities and make this value as large as possible. Since the log function is a monotonically increasing function, the problem can be cast in the SINRs. The objective can then be written as $\min_k w_k \text{SINR}_k$ where $w_k > 0$ are weighting factors to prioritize different users and enable us to achieve any point on the Pareto boundary of the achievable rate region (R_1, \dots, R_K) by varying the weights [27]. The weighted max-min fairness problem can be formulated as follows:

$$\begin{aligned} & \underset{\{p_k\}, t}{\text{maximize}} && t \\ & \text{subject to} && p_k \leq P_k, \forall k, \\ & && p_k \geq 0 \forall k, \\ & && w_k \text{SINR}_k \geq t, \forall k. \end{aligned} \tag{13}$$

With a fixed t , problem (13) can be solved using methods for solving problem (4) by omitting the power constraints. Then the optimal t can be found via bisection search such that the power constraints are satisfied. There are more advanced methods which can solve (13) efficiently in a distributed manner. For instance the Perron-Frobenius theory can be applied to solve (13) in the case that effect of noise is ignored [28]. Later generalization includes the approach using non-linear Perron-Frobenius theory [29] and the the Fast-Lipschitz optimization approach [30, 31].

2.3.2 Weighted Sum Performance

Another important and commonly used utility function is the weighted sum performance. This problem is proposed to maximize the total system throughput, and weights are included to provide some fairness between different users. We define the weighted sum SE by choosing $U(R_1, \dots, R_K) = \sum_{k=1}^K w_k R_k$. Power allocation that maximizes sum SE when interference is

present is known to be an NP-hard problem in general [12]. Therefore much work has focused on finding the local optimal solutions with low complexity. There also exists a line of work that optimizes the weighted sum with global optimization techniques [32–34]. However they are limited by the computational complexity and are hard to apply in practice, and can only act as a benchmark.

2.4 Joint Beamformer Design and Power Allocation

When the BS is equipped with multiple antennas, this creates more degrees of freedom for resource allocation. By transmit and receive beamforming, the signal of a particular user can be strengthened which brings array gain, meanwhile interference can be suppressed. Therefore the beamformer design can be very beneficial in increasing the users' data rates. To optimize the beamformer and power allocation simultaneously is therefore a problem that attracts a lot of interest [14, 15].

In the uplink, the optimal joint beamformer design and power allocation problem can be formulated as

$$\begin{aligned}
 & \underset{\{\mathbf{w}_k\}, \{p_k\}}{\text{maximize}} && U(R_1, \dots, R_K) \\
 & \text{subject to} && p_k \leq P_k, \forall k, \\
 & && p_k \geq 0 \forall k, \\
 & && \|\mathbf{w}_k\| \leq 1, \forall k
 \end{aligned} \tag{14}$$

where \mathbf{w}_k is the receive beamformer for user k and R_k is now defined as

$$R_k = \log_2 \left(1 + \frac{p_k |\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j \neq k} p_j |\mathbf{h}_j^H \mathbf{w}_k|^2 + \|\mathbf{w}_k\|^2} \right), \tag{15}$$

with \mathbf{h}_k denotes the channel between user k and the BS.

The optimal receive beamformer was shown to be the Wiener filter, also called receive minimum mean squared error (MMSE) filter as follows:

$$\mathbf{w}_k = \frac{(\mathbf{I}_M + \sum_i p_i \mathbf{h}_i \mathbf{h}_i^H)^{-1} \mathbf{h}_k}{\|(\mathbf{I}_M + \sum_i p_i \mathbf{h}_i \mathbf{h}_i^H)^{-1} \mathbf{h}_k\|}. \tag{16}$$

With a fixed beamformer, the power allocation problem becomes the same as in the previous section. However, this beamformer itself contains the power allocation parameters, therefore an iterative procedure is proposed [35] to solve the problem by first fixing the power allocation and optimize the beamformer, then update the power allocation with the known methods.

2.5 Receiver with Successive Interference Cancellation Capability

For the uplink scenario, the channel is modeled as a multiple access channel (MAC) in information theory. It is known that for MAC channel, performing successive interference cancellation (SIC) at the receiver can achieve any points in the capacity region [36]. This introduces another way of defining the SINR for user k as:

$$\text{SINR}_k^{\text{SIC}} = \frac{h_k p_k}{1 + \sum_{j>k} h_j p_j}. \quad (17)$$

Performing SIC at the receiver provides gain in the rate and the order of doing SIC does not matter. This statement holds true even when we extend to multiple antenna scenarios. Performing optimal received MMSE filtering together with SIC can achieve any points in the capacity region. Therefore in the uplink, SIC is theoretically optimal by introducing extra computational burden at the BS compared to treating interference as noise.

This idea has been borrowed to the downlink as well, by using super-position coding on the data for different users. Assuming SIC is used at the user with stronger channel, some gains in the rate are obtained. This idea is named non-orthogonal-multiple-access (NOMA) in the power domain [37] recently and attracts much research interest from both academia and industry [38–44]. The gain of NOMA in the single antenna case can be clearly seen. However when the BS is equipped with multiple antennas, it is not obvious whether it is worthwhile to perform SIC at the user side as the BS has the possibility to separate them by using beamforming techniques. Most literature in NOMA assumes perfect CSI when performing the analysis [39–41], and this leads to over optimistic conclusions that NOMA always provides gains. In paper E we will take a careful look at this with more practical assumptions that CSI

in both uplink and downlink is obtained through pilot training. A different conclusion is drawn and results there suggest that the marriage between massive MIMO and NOMA is probably not a good match.

Chapter 3

New Aspects of Power Allocation and Precoder Design in Massive MIMO

The use of massive MIMO at the BS introduces both challenges and opportunities in the transceiver design. One good piece of news is that linear processing techniques have been shown to be almost optimal. Two commonly used beamformers used in the massive MIMO literature are maximum ratio combining (MRC) and zero-forcing (ZF). They are shown to perform close to the optimal MMSE receiver in the low and high SNR regime respectively in single-cell systems. Another good piece of news is that an ergodic achievable rate can be found in closed form. The resource allocation can be much simplified as it only has to be redone when the large-scale fading changes, which is usually in the order of hundreds of coherence time. The large-scale fading is also approximately the same for different frequencies. This also reduces the control signal overhead of the system and the efforts on performing power allocation.

The bad news is that the channel estimation quality is critical in massive MIMO, therefore the commonly used perfect CSI assumptions are no longer valid. One has to take into account the channel estimation errors when optimizing the rate performance of the systems. This brings new challenges in the area of power allocation. Another bad news is that with large numbers of antennas at the BS, cheaper low-end hardware is preferred to keep the

cost down. When low-end hardware is used at the BS, we have to take into account the signal distortion due to the hardware impairment when performing power allocation.

3.1 The Role of CSI

As we saw from the previous sections, power allocation on the data has caught much attentions. Nevertheless, power allocation on the pilots is less exploited until recently. In massive MIMO systems, good enough channel estimation quality is needed for coherent beamforming. Now the problem is how to characterize the channel estimation quality?

With TDD operation, uplink pilots are sent for channel estimation. Consider a massive MIMO BS with M antennas serving K users, we denote the flat fading channel matrix between the BS and the users by $\mathbf{G} \in \mathbb{C}^{M \times K}$, where the k^{th} column \mathbf{g}_k represents the channel response to user k . Further, we assume that it has the distribution

$$\mathbf{g}_k \sim \text{CN}(\mathbf{0}, \beta_k \mathbf{I}), \quad k = 1, 2, \dots, K, \quad (18)$$

which is a circularly symmetric complex Gaussian random vector. The variance $\beta_k > 0$ represents the large-scale fading including path loss and shadowing, and is normalized by the noise variance at the BS to simplify the notation. The large-scale fading coefficients are assumed to be known at the BS as they are varying slowly (in the time scale of hundreds of coherence intervals) and can be easily estimated.

In each coherence interval, user k transmits its orthogonal pilot sequence with power p_p^k and length τ_p to enable channel estimation at the BS. We assume that minimum mean-squared error (MMSE) channel estimation is carried out at the BS to obtain the small-scale coefficients. This gives an MMSE estimate of the channel vector from user k as [45]

$$\hat{\mathbf{g}}_k = \frac{\sqrt{\tau_p p_p^k \beta_k}}{1 + \tau_p p_p^k \beta_k} \left(\sqrt{\tau_p p_p^k} \mathbf{g}_k + \mathbf{n}_p^k \right). \quad (19)$$

where $\mathbf{n}_p^k \sim \text{CN}(\mathbf{0}, \mathbf{I}_M)$ models the additive noise in the channel estimation process. In this case the variance of the elements in the estimate is denoted

by

$$\gamma_k = \frac{\tau_p p_p^k \beta_k^2}{1 + \tau_p p_p^k \beta_k}. \quad (20)$$

This quantity characterizes the channel estimation quality as $\beta_k - \gamma_k$ represents the variance of channel estimation error. Note that other channel estimators can be applied, however it will lead to a different channel estimation quality and is usually worse than applying the MMSE channel estimator.

3.2 Joint Pilot and Data Power Allocation

From (20) we observe that the channel estimation quality depends on the pilot power and the length of the training sequence which affect the data rate indirectly, meanwhile the data power p_d^k affects the data rate directly. To compare different power allocation schemes, a new form of power constraint should be formulated. We impose the following constraint on the total transmit energy over a coherence interval:

$$\tau_p p_p^k + (T - \tau_p) p_d^k \leq E_k, \quad k = 1, \dots, K \quad (21)$$

where E_k is the total energy budget for user k within one coherence interval. In this case, τ_p , p_p^k and p_d^k are optimization variables. Therefore the optimal power allocation should consider the case where each user can freely choose how to allocate its energy budget on the pilots and payload data. To compare the performance of different power allocation schemes, we take the achievable SE as our performance metric. When MRC is applied at the BS, an achievable SE can be written as [45]

$$R_k = \left(1 - \frac{\tau_p}{T}\right) \log_2(1 + \text{SINR}_k), \quad (22)$$

where

$$\text{SINR}_k^{\text{MRC}} = \frac{M p_d^k \gamma_k}{1 + \sum_{j=1}^K \beta_j p_d^j}. \quad (23)$$

Similarly when ZF is applied at the BS, only the SINR is changed and it can be written as [45]

$$\text{SINR}_k^{\text{ZF}} = \frac{(M - K) p_d^k \gamma_k}{1 + \sum_{j=1}^K p_d^j (\beta_j - \gamma_j)}. \quad (24)$$

The general power allocation problem is formulated as

$$\begin{aligned} & \underset{\tau_p, \{p_p^k\}, \{p_d^k\}}{\text{maximize}} && U(R_1, \dots, R_K) \\ & \text{subject to} && \tau_p p_p^k + (T - \tau_p) p_d^k \leq E_k, \quad \forall k, \\ & && p_p^k \geq 0, p_d^k \geq 0, \quad \forall k, \\ & && K \leq \tau_p \leq T. \end{aligned} \tag{25}$$

We see that the methods developed in the literature [] does not apply to this new problem formulation. Therefore a new framework for power allocation has to be developed. In Paper A, we develop a framework for joint pilot and data power allocation scheme based on the discussions above. As this framework is relatively new, many interesting problems are still open. For example Paper A also discusses the case when the channel model is changed.

3.3 Help of Asymptotic Analysis

Despite the advantages brought by massive MIMO, it also introduces new challenges in designing the system. In single-cell systems, linear precoders like ZF and MRC are asymptotically optimal in high and low SNR regime, this only holds true when the number of antenna M is much greater than the number of users K . However with the increasing demand of the number of active devices, this conclusion does not always hold. Therefore we see the need of designing optimal precoders to cope with these scenarios.

A general problem of designing optimal precoders is the computational complexity. As M and K are both large, this introduces tremendous computational burden and therefore the methods for optimizing the precoders in the literature are not feasible in massive MIMO. Luckily, we can apply results in random matrix theory [46–49] to simplify the problem. When M and K are both large, the randomness of the small-scale fading can be averaged out and the only parameters left are the large-scale fading coefficients. This is known as the deterministic equivalent in the random matrix theory literature.

Making use of the deterministic equivalent expressions, we can simplify the design problems in large dimensions. This greatly reduces the computational burden at the BS. In Paper B we will showcase the application of this idea and we will see that problems that are originally hard, for example the sum rate maximization problem, can be solved under mild conditions in large

dimensions. Moreover we will also show that the computational complexity is almost the same as ZF which is practically realizable in real time [50].

3.4 Power Allocation with Imperfect Amplifiers

While the previous parts are devoted to power allocation in the uplink, we focus on the downlink in this subsection. More specifically, we discuss how imperfect non-linear amplifiers change the power allocation problem. The sum radiated power constraint considered in the literature is not realistic as it does not take into account the maximum output constraint of an individual power amplifier (PA). Nevertheless, in many applications the power consumed by the power amplifiers consists of both the output power and the power losses in the hardware. As low-end power amplifiers are foreseen to be used at massive MIMO BSs, it is essential to take the PA into account when designing the transmitters at base stations.

For the consumed power we use the amplifier modeling in [51, 52], and set

$$\frac{p_i}{q_i} = \eta_{\max} \left(\frac{p_i}{P_{\max}} \right)^{1-\epsilon}, \quad (26)$$

where ϵ is a parameter with $\epsilon \in [0, 0.5]$, p_i is the radiated power on antenna i , P_{\max} is the maximum allowed radiated power, q_i is the consumed power on antenna i , and $\eta_{\max} \in [0, 1]$ is the maximum power efficiency obtained when $p_i = P_{\max}$. If $\eta_{\max} = 0$, all power is dissipated. The maximum power efficiency is a fixed parameter. It can be common or different for all the employed PAs at the transmitter, but we assume it to be the same for all the PAs as it does not change much in practice. From the above, we can write

$$q_i = \frac{1}{\eta_{\max}} p_i^\epsilon P_{\max}^{1-\epsilon}. \quad (27)$$

The consumed power q_i is thus proportional to the ϵ -th power of the radiated power and it will always be greater than the radiated power. When $\epsilon = 0$ the consumed power will be the same whenever the amplifier is on regardless of the radiated power.

We can see that this amplifier model changes the power constraints in the power allocation problems. This non-convex constraint is hard to deal with when we are optimizing the power allocation, In Paper C and Paper D we will look at different scenarios where we apply this model.

Chapter 4

Background on Optimization

4.1 General Optimization Problems

Many engineering problems will be eventually converted into an optimization problem. In particular power allocation problems are usually cast as optimization problems where we optimize the power allocation parameters to achieve a certain goal. In this thesis, we consider the problem to maximize certain objective functions or minimizing the power consumption while satisfying some predefined constraints. The problems we have can be formulated in the following form:

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} && f(\mathbf{x}) \\ & \text{subject to} && g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \\ & && h_j(\mathbf{x}) = 0, \quad j = 1, \dots, p. \end{aligned} \tag{28}$$

This is the general form of optimization problems. If we are having a maximization problem, we can always change the problem to a minimization problem by taking the negative sign of the function we are maximizing. Here are some terminologies we will use:

- optimization variables, which usually are vectors $\mathbf{x} \in \mathbb{R}^n$, but can also be matrices in some cases
- objective function $f(\mathbf{x})$ from \mathbb{R}^n to \mathbb{R}

- inequality constraints $g_i(\mathbf{x}) \leq 0$, and there are m of these
- equality constraints $h_j(\mathbf{x}) = 0$, and there are p of these
- feasible points: all \mathbf{x} that satisfy all the inequality constraints and equality constraints
- feasible set: the set of all feasible points
- global optimal solution: feasible point \mathbf{x}^* that achieves the minimum (or maximum) value among all the feasible points
- local optimal solution: feasible point \mathbf{x}^* that achieves the minimum (or maximum) value in a neighbourhood around \mathbf{x}^* , usually characterized by the Karush-Kuhn-Tucker (KKT) conditions [53]

In general, optimization problems are not easy to solve, finding the global optimal solutions has high computational complexity. Therefore in many cases people are satisfied with finding a local optimal solution by using some heuristics. Nevertheless, for some classes of problems we can actually find the global optimal solutions efficiently. Among them convex optimization problem is the most important class. For convex optimization problems, every local optimum is globally optimal. Therefore the theory is well understood and many efficient algorithms are developed for general convex optimization problems. A problem is called a convex problem if the feasible set is a convex set, i.e. $g_i(\mathbf{x})$ are convex functions and $h_j(\mathbf{x})$ are affine functions, and the minimization (maximization) is over a convex (concave) objective function $f(x)$. For more details about convex optimization, readers are referred to the classic textbook [53].

In the following we will review some useful techniques used to transform a problem to convex form. After that, an algorithm for finding local optimal solutions is presented for general optimization problems.

4.2 Some Convexity Preserving Transformation

There are many transformations that can preserve the convexity of a problem. Here we introduce some of them that are used in the included papers in this thesis. We hope this can help the readers to understand the techniques applied in the included papers.

4.2.1 Epigraph

The problem (28) can be equivalently written in the following epigraph form:

$$\begin{aligned}
 & \underset{\mathbf{x}, t}{\text{minimize}} && t \\
 & \text{subject to} && f(\mathbf{x}) \leq t \\
 & && g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \\
 & && h_j(\mathbf{x}) = 0, \quad j = 1, \dots, p.
 \end{aligned} \tag{29}$$

The epigraph form will always give the same objective values as the original problem and the optimal solution from \mathbf{x}, t is also the optimal solution for the original problem. This follows immediately from the fact that at the optimal point the constraint $f(\mathbf{x}) \leq t$ will be active, i.e. $f(\mathbf{x}) = t$. Otherwise we can always decrease t to make the inequality active and at the same time yield a lower objective value. This method is particularly useful when we are solving some problems with non-smooth objective function, e.g. the max-min problem.

4.2.2 Variable Substitution

Another common way of transforming a problem is using variable substitution. This can sometimes exploit the hidden convex structure in the original problem formulation.

One example is the geometric programming (GP) problem:

$$\begin{aligned}
 & \underset{\mathbf{x}, t}{\text{minimize}} && f(\mathbf{x}) \\
 & \text{subject to} && g_i(\mathbf{x}) \leq 1, \quad i = 1, \dots, m, \\
 & && h_j(\mathbf{x}) = 1, \quad j = 1, \dots, p.
 \end{aligned} \tag{30}$$

Here $h_j(\mathbf{x})$ is in the form $h_j(\mathbf{x}) = h_j(x_1, \dots, x_n) = c_j \prod_n x_n^{a_{jn}}$ with $c_j > 0$, $a_{jn} \in \mathbb{R}$, which is defined as a monomial. $f(\mathbf{x})$ and $g_i(\mathbf{x})$ are positive sums of monomials, defined as the posynomial. This problem formulation is obviously non-convex, however with variable substitution it can be transformed to a convex problem. Introduce the change of variable $y_n = \log(x_n)$ for

every variable x_n . Monomials become exponentials of affine functions and posynomials become sum of exponentials of affine functions. Finally taking the logarithm of every objective and constraint function, we obtain a convex problem as the log-sum-exponential functions are convex. The equivalent convex formulation is as follows:

$$\begin{aligned} & \underset{\mathbf{y}}{\text{minimize}} && \log \left(\sum e^{\sum_n a_n y_n} \right) \\ & \text{subject to} && \log \left(\sum_i e^{\sum_n a_{in} y_n} \right) + \log(c_i) \leq 0, \quad i = 1, \dots, m, \\ & && \sum_n a_{jn} y_n + \log(c_j) = 0, \quad j = 1, \dots, p. \end{aligned} \tag{31}$$

4.2.3 Perspective Transformation

The perspective transformation is not a commonly used technique, however it is applied in Paper A. Here we introduce the basic definition and the properties of the perspective function. For any function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the perspective of f is defined as

$$g(\mathbf{x}, t) = t f(\mathbf{x}/t), \tag{32}$$

where $g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ and the domain of g (the set where function g is defined) is $\mathbf{dom} g = \{(\mathbf{x}, t) | \mathbf{x}/t \in \mathbf{dom} f, t > 0\}$.

Examples of perspective functions include

- the perspective of a linear function is itself,
- the perspective of $\mathbf{x}^T \mathbf{x}$ is $\frac{\mathbf{x}^T \mathbf{x}}{t}$.
- the perspective of $\log(x)$ is $t \log\left(\frac{x}{t}\right)$

One important property of the perspective function is that it preserve convexity, i.e. f is convex if and only if g is convex. This property will be used in Paper A.

4.3 General Sufficient Conditions for Global Optimality

In this section, we introduce a general sufficient condition for global optimality that holds for any optimization problem. This condition is less well known compared to the famous KKT conditions which are necessary conditions for any local optimal point under some regularity conditions. In some occasions this sufficient condition is confused with the KKT conditions. Therefore we provide the general sufficient conditions here and distinguish the differences from the KKT conditions.

For the need of discussion, we define the Lagrange multiplier as $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, associated with the inequality and equality constraints respectively, then the Lagrangian function of (28) is

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \mu_j h_j(\mathbf{x}). \quad (33)$$

We are ready to present the optimality conditions, which is a slightly more general version than the ones appearing in [54, 55] to include the equality constraint:

Lemma 1. The vector \mathbf{x}^* is a global optimal solution to (28) if \mathbf{x}^* is feasible and there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that the following conditions are satisfied:

$$\mathbf{x}^* \in \arg \min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \quad (34)$$

$$\lambda_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, \dots, m. \quad (35)$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, m. \quad (36)$$

Proof. For any feasible \mathbf{x} , we have the following chain of inequalities

$$f(\mathbf{x}^*) = L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \leq L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \leq f(\mathbf{x}). \quad (37)$$

The equality is due to the condition in (35) and the fact that \mathbf{x}^* is feasible in problem (28), i.e., $g_i(\mathbf{x}^*) \leq 0, \forall i$ and $h_j(\mathbf{x}^*) = 0, \forall j$. The first inequality holds as \mathbf{x}^* is the minimizer of the Lagrangian function. The last inequality holds as the Lagrangian is always a lower bound on the original objective

function for any feasible \mathbf{x} , since $\lambda_i^* g_i(\mathbf{x}) \leq 0$ and $\mu_j^* h_j(\mathbf{x}) = 0$ for any feasible \mathbf{x} . Optimality follows since \mathbf{x}^* is feasible and the resulting objective function is less than or equal to any other feasible point. \square

This optimality result is a general sufficient condition, but the major drawback is that it is not necessary for all global optimal solution. There are some cases that this condition does not hold, where $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ satisfying the conditions in Lemma 1 do not exist. Moreover, solving the optimization problem in (34) itself can be tough when the original problem is non-convex. Despite the difficulties of utilizing the results, it can be useful in finding globally optimal solutions for non-convex problems or serves as certificate for reaching global optimality. In Paper B we will show an example of applying Lemma 1 to solve a non-convex problem.

The sufficient condition in Lemma 1 is in some form similar to the famous KKT conditions which, under some regularity conditions, described a set of necessary conditions for any local optimal solutions. The KKT conditions can be written as,

$$\mathbf{0} \in \partial_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \quad (38)$$

$$\lambda_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, \dots, m, \quad (39)$$

$$\lambda_i^* \geq 0, \quad i = 1, \dots, m, \quad (40)$$

$$g_i(\mathbf{x}^*) \leq 0, h_j(\mathbf{x}^*) = 0, \quad \forall i, j \quad (41)$$

where $\partial_{\mathbf{x}}$ represents the sub-gradient with respect to \mathbf{x} . Comparing the sufficient conditions and the KKT conditions, the difference appears in (34). We can see that the conditions in Lemma 1 implies the KKT conditions, but not vice versa. We can expect this to happen as a globally optimal solution must be a locally optimal solution. Therefore this suggests a general strategy by enumerating all the points that satisfy the KKT conditions, and the conditions in Lemma 1 can be used as a certificate for achieving global optimality.

When the problem itself is a convex problem, i.e. when $f(\mathbf{x}), g_i(\mathbf{x}), \forall i$ are convex functions and $h_j(\mathbf{x}), \forall j$ are affine functions, then conditions in Lemma 1 is equivalent to the KKT conditions, and therefore are both sufficient and necessary when some regularity conditions are satisfied.

4.4 Successive Convex Optimization

In many occasions, the optimization we formulate is non-convex and has no hidden convex structure. In this case finding the global optimal solution is difficult. Global optimization techniques can be applied but the complexity grows exponentially with the number of variables and constraints. In this case, algorithms with lower complexity that find local optimal solutions is an option. The successive convex approximation is an algorithmic framework [56] that can be used to solve general non-convex problems and give a local optimal solution. In power allocation applications, this framework was used to handle non-convex constraints [13]. There exist other terminologies with the same idea as this framework, for instance ‘Majorization Minimization’ and ‘Inner Approximation’. The idea is to solve a series of approximated problems where the non-convex constraint is approximated with a convex constraint in each problem. This method can be applied to a problem with any finite number of non-convex constraints, here we illustrate the idea and the procedure for the case of only one non-convex constraint.

Consider a non-convex optimization problem

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && f(\mathbf{x}) \\ & \text{subject to} && g(\mathbf{x}) \leq 0 \\ & && x \in \mathcal{C} \end{aligned} \tag{42}$$

where \mathcal{C} is an arbitrary convex set and function $g(\mathbf{x})$ is non-convex. Note that if the objective function $f(x)$ is non-convex, we can always move it to the constraints by using the epigraph form. We approximate the non-convex function $g(\mathbf{x})$ with a convex function $g_k(\mathbf{x})$ in the k -th iteration. The convex optimization problem to be solved in the k -th iteration is

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} && f(\mathbf{x}) \\ & \text{subject to} && g_k(\mathbf{x}) \leq 0 \\ & && x \in \mathcal{C}. \end{aligned} \tag{43}$$

If we construct a family of functions $g_k(\mathbf{x})$ in each iteration k satisfying the conditions

1. $g(\mathbf{x}) \leq g_k(\mathbf{x}), \forall \mathbf{x} \in \mathcal{C}$,

2. $g(\mathbf{x}^{(k-1)}) = g_k(\mathbf{x}^{(k-1)})$, where $\mathbf{x}^{(k-1)}$ is the solution from the previous iteration,
3. $\nabla g(\mathbf{x}^{(k-1)}) = \nabla g_k(\mathbf{x}^{(k-1)})$,

the algorithm will give a solution satisfying the KKT conditions for the original problem (42).

Outline of proof: The first condition is to ensure that the solution we get is feasible for the original problem. The second condition ensures that the solution from the previous iteration is feasible for the current iteration. As a result the objective value of the original problem decreases after every iteration since the solution from the previous iteration is a feasible solution to the problem in the current iteration. The second and third conditions together guarantee that the KKT conditions for the original problem will be satisfied at convergence.

To conclude, we can obtain a suboptimal solution to (42) with the procedure described in Algorithm 1.

Algorithm 1 Successive convex optimization for problem (42)

- 1: choose $\mathbf{x}^{(0)}$ satisfying the constraints and initialize $k = 1$
 - 2: **repeat**
 - 3: form the k^{th} approximated problem (43) of (42) by approximating $g(\mathbf{x})$ with $g_k(\mathbf{x})$ around $\mathbf{x}^{(k-1)}$
 - 4: solve the k^{th} approximated problem to obtain $\mathbf{x}^{(k)}$,
 - 5: $k \leftarrow k + 1$
 - 6: **until** convergence
 - 7: **return** $\mathbf{x}^{(k)}$
-

The above Algorithm 1 can handle the non-convex constraints. On the contrary, when the objective function is non-convex, a recently developed framework [57] can be applied. In each iteration the objective function is approximated with a pseudo-convex function (a larger class of functions compared to the class of convex functions). The approximated function need not be upper or lower bound of the original function, however a proper line search has to be done between different iterations.

Chapter 5

Future Work

Massive MIMO has undergone a rapid development in the last ten years, and the theoretic results are mature that the potentials and limitations are mostly known. This thesis provides some new directions to further optimize the performance in massive MIMO. These analyses and methods are done for single-cell massive MIMO, and the extension to multi-cell massive MIMO would be interesting. When extending to multi-cell scenarios, interference from other cells and more importantly, the pilot contamination effects, needs to be accounted for when optimizing the SE performance. For example, one can extend the joint pilot and data power allocation framework to multi-cell case and see if pilot contamination effects could be reduced.

As the manufacturing techniques are maturing and the communication frequency bands increases, there will be more antennas implemented in mobile devices. Therefore another interesting direction is to look at the potentials of utilizing multiple antennas at the terminals for beamforming to enhance both the channel estimation quality and the SE performance. There are some preliminary results in the literature [58–60], and it is worth digging deeper in this topic.

A new promising wireless technology after massive MIMO is the cell free massive MIMO (or distributed massive MIMO) [61,62]. The idea is to distribute the antennas in a large area and make use of all antennas for joint-processing. Extension of the framework in Paper A and Paper B could be of interest. As the antennas are not collocated, a distributed processing scheme that does not requires communication overhead between the distributed antennas is

worth investigating.

In Paper B we make use of large-system analysis to reduce the computational complexity of precoder design problem. From this example we see that in some cases asymptotic expressions in large-system could be useful for some intermediate steps in solving optimization problems. This motivates the extension of this idea to other applications apart from the communication area. For instance, in data science, we are dealing with a large amount of data samples and unknown parameters. It would be interesting to look into the possibility of applying large-system analysis in reducing the complexity of different algorithms.

Bibliography

- [1] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020 White Paper”, Jan. 2016.
- [2] T. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014
- [4] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: benefits and challenges,” *IEEE Journal of Selected Topics in Signal Processing*, vol.8, no.5, pp. 742–758, Oct. 2014
- [6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Transactions on Communications*, vol.61, no.4, pp.1436–1449, Apr. 2013.
- [7] H. Yang and T. L. Marzetta, “Capacity performance of multicell large-scale antenna systems,” in *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2013.
- [8] J. Hoydis, S. ten Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

- [9] H. Yang, and T. Marzetta, “Performance of conjugate and zero-forcing beamforming in large-scale antenna systems,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [10] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, “Argos: practical many-antenna base stations,” in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, pp. 53–64, Aug. 2012.
- [11] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, “A flexible 100-antenna testbed for massive MIMO,” in *Proceedings of the IEEE Globecom Workshops*, pp. 287–293, Dec. 2014.
- [12] Z. -Q. Luo and S. Zhang, “Dynamic spectrum management: Complexity and duality,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [13] M. Chiang, C. W. Tan, D. Palomar, D. O’Neill, and D. Julian, “Power control by geometric programming,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.
- [14] S. Shi, M. Schubert, and H. Boche, “Rate optimization for multiuser MIMO systems with linear processing,” *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.
- [15] E. Björnson and E. Jorswieck, “Optimal resource allocation in coordinated multi-cell systems,” *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.
- [16] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, “Massive MIMO systems with non-ideal hardware: energy efficiency, estimation, and capacity limits,” *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.
- [17] H. V. Cheng, E. Björnson, and E. G. Larsson, “Optimal pilot and payload power control in single-cell massive MIMO systems,” *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2363–2378, May 2017.
- [18] H. V. Cheng, E. Björnson, and E. G. Larsson, “Uplink pilot and data power control for single cell massive MIMO systems with MRC,” in *Proceedings of International Symposium on Wireless Communication Systems (ISWCS)*, 2015, pp. 396–400.

-
- [19] H. V. Cheng, E. Björnson, and E. G. Larsson, “Semi-closed form solution for sum rate maximization in downlink multiuser MIMO via large-system analysis,” in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2018.
- [20] H. V. Cheng, D. Persson, and E. G. Larsson, “MIMO capacity under power amplifiers consumed power and per-antenna radiated power constraints,” in *Proc. IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2014, pp. 179–183.
- [21] H. V. Cheng, D. Persson, E. Björnson, and E. G. Larsson, “Massive MIMO at night: On the operation of massive MIMO in low traffic scenarios,” in *Proceedings of 2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 1697–1702.
- [22] H. V. Cheng, E. Björnson and E. G. Larsson, “Performance analysis of NOMA in training-based multiuser MIMO systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 372–385, Jan. 2018.
- [23] H. V. Cheng, E. Björnson and E. G. Larsson, “NOMA in multiuser MIMO systems with imperfect CSI,” in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–5.
- [24] F. Bock and B. Ebstein, “Assignment of transmitter powers by linear programming,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 6, no. 2, pp. 36–44, Jul. 1964.
- [25] G. J. Foschini and Z. Miljanic, “A simple distributed autonomous power control algorithm and its convergence,” *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 641–646, Nov. 1993.
- [26] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
- [27] E. Björnson, E. Jorswieck, M. Debbah, and B. Ottersten, “Multi-objective signal processing optimization: the way to balance conflicting metrics in 5G systems,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 14–23, Nov. 2014.

- [28] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, pp. 57–62, Feb. 1992.
- [29] C. W. Tan, M. Chiang, and R. Srikant, "Maximizing sum rate and minimizing MSE on multiuser downlink: optimality, fast algorithms and equivalence via max-min SINR," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6127–6143, Dec. 2011.
- [30] C. Fischione, "Fast-Lipschitz optimization with wireless sensor networks applications," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2319–2331, Oct. 2011.
- [31] C. Fischione and M. Jakobsson, "Optimality of radio power control via Fast-Lipschitz optimization," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2589–2601, Jun. 2016.
- [32] L. P. Qian, Y. J. Zhang, and J. Huang, "MAPEL: achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.
- [33] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [34] P. C. Weeraddana, M. Codreanu, M. Latva-aho, A. Ephremides, and C. Fischione, "A review of weighted sum-rate maximization in wireless networks," *Foundations and Trends in Networking*, Vol. 6, No 1–2, pp. 1–163, 2012.
- [35] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [36] A. E. Gamal and Y. H. Kim, *Network Information Theory*, Cambridge University Press, Cambridge, U. K., 2012.
- [37] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proceedings of 2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, Jun. 2013, pp. 1–5.

-
- [38] L. Dai, B. Wang, Y. Yuan, S. Han, C. L. I, and Z. Wang, “Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, Sep. 2015.
- [39] Z. Ding, F. Adachi, and H. V. Poor, “The application of MIMO to non-orthogonal multiple access,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 537–552, Jan. 2016.
- [40] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, “On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users,” *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
- [41] Q. Sun, S. Han, C. L. I, and Z. Pan, “On the ergodic capacity of MIMO NOMA systems,” *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 405–408, Aug. 2015.
- [42] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C. L. I, and H. V. Poor, “Application of non-orthogonal multiple access in LTE and 5G networks,” *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [43] V. W. Wong, R. Schober, D. W. K. Ng, and L. Wang, *Key Technologies for 5G wireless systems*. Cambridge, U. K.: Cambridge University Press, 2017.
- [44] 3GPP TD RP-150496, “Study on downlink multiuser superposition transmission.”
- [45] T. Marzetta, E. G. Larsson, H. Yang, and H. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge University Press, 2016.
- [46] R. Zakhour and S. V. Hanly, “Base station cooperation on the downlink: large system analysis,” *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.
- [47] Y. Huang, C. W. Tan, and B. D. Rao, “Large system analysis of power minimization in multiuser MISO downlink with transmit-side channel correlation,” in *Proceedings of 2012 International Symposium on Information Theory and its Applications*, Oct. 2012, pp. 240–244.
- [48] L. Sanguinetti, A. L. Moustakas, E. Björnson, and M. Debbah, “Large system analysis of the energy consumption distribution in multi-user

- MIMO systems with mobility,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1730–1745, Mar. 2015.
- [49] L. Sanguinetti, R. Couillet, and M. Debbah, “Large system analysis of base station cooperation for power minimization,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5480–5496, Aug. 2016.
- [50] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, “The world’s first real-time testbed for massive MIMO: Design, implementation, and validation,” *IEEE Access*, vol. 5, pp. 9073–9088, 2017.
- [51] S. Mikami, T. Takeuchi, H. Kawaguchi, C. Ohta, and M. Yoshimoto “An efficiency degradation model of power amplifier and the impact against transmission power control for wireless sensor networks,” in *Proceedings of IEEE Radio and Wireless Symposium*, Jan. 2007, pp. 447-450.
- [52] D. Persson, T. Eriksson, and E. G. Larsson, “Amplifier-aware multiple-input multiple-output power allocation,” *IEEE Communications Letters*, vol. 17, no. 6, pp. 1112–1115, Jun. 2013.
- [53] S. Boyd, L. Vandenberghe, “*Convex Optimization*”, Cambridge University Press, 2004
- [54] D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 3rd Edition 2016.
- [55] D.P. Bertsekas, A. Nedić, and A.E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific optimization and computation series. Athena Scientific, 2003.
- [56] B. R. Marks and G. P. Wright, “A general inner approximation algorithm for nonconvex mathematical programs,” *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.
- [57] Y. Yang and M. Pesavento, “A unified successive pseudoconvex approximation framework,” *IEEE Transactions on Signal Processing*, vol. 65, no. 13, pp. 3313–3328, Jul. 2017.
- [58] N. N. Moghadam, H. Shokri-Ghadikolaei, G. Fodor, M. Bengtsson and C. Fischione, “Pilot precoding and combining in multiuser MIMO networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1632–1648, Jul. 2017.

- [59] X. Li, E. Björnson, S. Zhou and J. Wang, “Massive MIMO with multi-antenna users: When are additional user antennas beneficial?”, in *Proceedings of 2016 23rd International Conference on Telecommunications (ICT)*, 2016, pp. 1–6.
- [60] E. L. Bengtsson and P. C. Karlsson and F. Tufvesson and J. Vieira and S. Malkowsky and L. Liu and F. Rusek and O. Edfors, “Transmission schemes for multiple antenna terminals in real massive MIMO systems,” in *Proceedings of 2016 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016, pp. 1–6.
- [61] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson and T. L. Marzetta, “Cell-free massive MIMO versus small cells,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [62] H. Yin, D. Gesbert and L. Cottatellucci, “Dealing with interference in distributed large-scale MIMO systems: a statistical approach,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 942–953, Oct. 2014.

Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-145674>